

Index

Page numbers followed by *f* indicate figures; page numbers followed by *t* indicate tables.

- Adaptive rejection metropolis sampling (ARMS), 98
- Adaptive shrinkage, 132
- Advanced Photo System (APS), 255
- Aggregation
 - levels of, 103
- Alternative information sets
 - use of, 179
- Assortment
 - utility of, 224*t*, 223–224
- Attainable utility, 123
- Autocorrelation function (ACF), 55
- Autoregressive discrete choice model
 - Bayesian spatial, 225
- Autoregressive matrix
 - elements of, 228

- Bayes, 15–16
- Bayes binary logit R code
 - hierarchical, 312–315
- Bayes estimators, 15
 - absolute error, 17
 - admissible, 18
 - decision theory of, 17–19
 - duality, 19
 - hyperparameter vector, 17
 - loss function, 17
 - parameter estimation, 17
 - risk of, 17–19
 - sampling properties of, 17–19
 - shrinkage, 25
 - squared error, 17
- Bayes factors
 - asymptotic methods for computation of, 163–165
 - computation of, 173–177
 - using importance sampling, 165–166
 - in conjugate setting, 162–163
 - Savage–Dickey density ratio, 162
 - using Markov chain Monte Carlo (MCMC) draws, 166–169
 - Newton–Raftery estimator, 176*f*
- Bayes linear regression
 - analysis of, 76
- Bayes model
 - hierarchical, 132
- Bayes modeling
 - hierarchical, 303–321
 - analysis, organizing data for, 305–312
 - bayesm*, 279
 - data description, 304*t*
 - demographic observations in, 308*t*
 - observations, 306*t*
 - in R, 279–321
 - R package, 279
- Bayes’ theorem, 13–15
 - Lebesgue measure, 14
 - posterior distribution, 14
- Bayesian
 - inference procedures, 15
- Bayesian analog, 16

- Bayesian approach
 - benefits of, 4–6
 - costs of, 4–6
 - nontrivial, 4
 - decision-making
 - integration of, 4
 - empirical, 70
 - to instrumental variables, 185–194
 - objectivity, 5
- Bayesian decision theory, 159–184
 - example of, 180–184
- Bayesian decision-maker, 160
- Bayesian essentials, 9–47
 - asymptotic methods, 35–36
 - conditioning of, 15
 - conjugacy, 20–21
 - distribution theory
 - concepts from, 9–13
 - exponential families, 20–21
 - identification of, 19–20
 - importance sampling, 37–40
 - integration methods, 35–36
 - likelihood principle, 15
 - multivariate analysis
 - examples of, 21–35
 - posterior, summarizing of, 16–17
 - prediction, 15–16
 - predictive distribution, 16
 - regression, 21–35
 - sufficiency, 20–21
- Bayesian estimation, 198–200
- Bayesian factors
 - Newton–Raftery (NR) estimate, 168
- Bayesian inference, 19–20, 159
 - computational phase of, 49
- Bayesian information criterion (BIC), 165
- Bayesian methods, 1, 155
- Bayesian model
 - archetypal data augmentation model, 171
 - bridge sampling methods, 169–170
 - Chib’s approach, 176
 - Chib’s method, 171–172
 - choice of, 159–184
 - Gelfand–Dey estimator, 170
 - selection of, 160–162
- Bayesian problems
 - Dirichlet distribution, 45
 - Gamma generation, 41
 - multinomial distribution, 44–45
 - multivariate normal distribution, 42–43
 - normal generation, 41
 - simulation primer for, 41–45
 - student t distribution, 42–43
 - truncated distributions, 41–42
 - uniform generation, 41
 - Wishart distribution, 43–44
 - inverted, 43–44
- bayesm* 279, 323
 - adapting codes, 327
 - extending codes, 327
 - installation of, 323–327
 - installing of, 284, 284*f*
 - key utilities, 332*t*
 - manual for, 326
 - MCMC functions, 328*t*
 - obtaining help on, 324–327
 - update of, 327
 - use of, 323–327
 - tips for, 324
- Bernoulli random variable, 9
- β
 - correlation with latents, 113*f*
- Beta
 - correlations of, 141*t*
 - standard deviations of, 141*t*
- Binary logit model
 - hierarchical, 279
- Binary probit, 104
 - Gibbs sampler for, 77, 77*f*, 107
- Bivariate density function, 10
- Bivariate distribution, 11*f*
- Brands
 - demand for, 207
- Bridge sampling methods, 169–170
- Camera
 - attributes of, 257*t*
 - levels of, 257*t*
- Chib’s approach, 176
 - densities, values of, 178*f*
- Chib’s method, 171–172
- Cholesky root, 6, 52
- Cobb–Douglas utility function, 209, 219
- Coefficient estimates
 - aggregate, 221*t*
 - posterior mean of, 233*t*
- Compensatory model, 256
- Comprehensive R Archive Network (CRAN), 8, 279
- Conditional distribution, 12
- Conditional modeling approach, 203–206
 - SAMP, 203
- Conjugacy, 20–21

- Conjugate priors
 - conditionally conjugate, 34
 - limitations of, 34–35
 - natural, 34
- Conjunctive model
 - part-worth estimates for, 264*t*
 - threshold estimates for, 263*f*, 266*t*
- Conjunctive screening, 264–266
 - background of, 253–254
 - choice set, 253
 - data for, 255–258
 - model for, 254–255, 259*t*
 - R implementation of, 266–267
 - results of, 259–264
 - rules for, 253–267
 - use of, 260
- Consumer packaged goods (CPG), 72
- Consumer purchase behavior
 - quantitative models of, 225
- Continuous random variable, 10
- Coors Light, 214
- Corner solution, 272
 - proof of, 213–214
- Correlation coefficient
 - prior distribution of, 121*f*
- Covariance matrices
 - Bayesian inference for, 28–30
 - trace operator, 29
- CPU speed, 7
- CRAN network, 323
- Cumulative distribution function (cdf), 41
- Customer satisfaction
 - survey questionnaire, 238*f*
- Cut-off parameter
 - role of, 244*f*
- Cut-offs, 259*f*

- Data augmentation, 75–79, 104, 198, 246
- Data frame, 286
- Debugging functions, 298–299
- Δ
 - posterior distribution of, 142*t*
- Demand model, 196–197
- Demand shifter, 195
- Demand systems, 274
- Demand theory, 122–128
- Density function, 10
- Detailing allocation
 - application to, 202
- Directed acyclic graph (DAG), 67
- Dirichlet distribution, 45
- Disaggregate information
 - valuation of, 179–180
- Discrete approximation, 59
- Discrete choice models, 122–128
- Discrete demand, 103–128
- Discrete dependent variables, 104
- Discrete distribution
 - joint, 241
- Discrete quantities
 - demand for, 124–125
- Discrete random variable, 9
- Distribution theory
 - Bernoulli random variable, 9
 - bivariate density function, 10
 - concepts from, 9–13
 - conditional distribution, 12
 - continuous random variable, 10
 - density function, 10
 - discrete random variable, 9
 - joint distribution, 10
 - marginal density, 11
 - number, expectation of, 10
 - probability mass function, 10
- Draw sequences
 - comparison of, 138*f*

- Endogeneity, 195–200
- Endogeneity bias, 185–186
- Endogeneity case
 - weak, MCMC draws, 195*f*
- Ergodicity, 56*f*, 57
- Expected revenues
 - posterior distribution of, 184*f*
- Extended sampler, 102

- Finite mixtures, 79
- Finite-mixture model, 144
- First-stage prior
 - diagnostic checks of, 155–156
- Forward chain
 - Markov property of, 61
- Fubini's theorem, 18

- Γ
 - posterior mean of, 321*t*
- Gamma distribution
 - inverse, 23
- Gelfand–Dey estimator, 170
- GHK algorithm, 40
- Gibbs sampler, 35, 63–65, 133, 145–148
 - bivariate normal, 52–57
 - autocorrelation function, 56*f*

- Gibbs sampler, (*continued*)
 ergodicity, 56*f*
 functioning of, 54*f*
 intermediate moves, 55*f*
 one thousand draws from, 56*f*
 convergence of, 64
 for seemingly unrelated regression (SUR)
 model, 65–67
 transition mechanism in, 52
 unconstrained
 performance of, 84–85
- Gibbs-style chain, 97
- GLS estimator, 67
- Group profiling, 235*t*
- Heterogeneity, 129–131, 271
 correlation analysis of, 249
 household, 220, 277
 normal model of, 154–155
 overcoming scale usage, 237–252
 background of, 237–240
 data, 247*f*, 246–251, 250*t*
 MCMC algorithm, 244–246
 model of, 240–244
 priors, 244–246
 R implementation, 252
 parametric distributions of, 156
 scale usage, 247–249
- Heterogeneous units
 hierarchical models for, 129–158
 adaptive shrinkage, 132
 Bayes model, 132
 heterogeneity, 130–131
 random effects model, 133
- Heteroskedasticity, 126, 146
- Hierarchical linear models, 70–75
 Gibbs sampler, 72*f*
- Hierarchical models
 inference for, 133–136
 panel data structures, 133
- hist command
 use of, 282*f*
- Hit frequency, 259
- Household coefficient
 posterior means of, 143*f*, 152*f*, 154*f*
- Household expenditure
 posterior estimates of, 222*f*
- Household purchase information
 valuing of, 180–184
- Hypothesis testing
 Neyman–Pearson approach, 161
- Importance function, 38
 fat-tailed, 39
- Importance sampling, 37–40, 89, 165
 weight distribution of, 177*f*
- In-sample fit, 231
- Independence metropolis chain, 89–90
- Instrumental variables, 185
 brushing technique, 187
 endogenous, 186
 Gibbs sampler, 189
 illustration of, 187*f*
 reduced form, 187
 strong
 MCMC draws, 191*f*
 weak, 193
 marginal posteriors for, 194*f*
 MCMC draws, 192, 193*f*
- Interdependent consumer preferences, 235*f*,
 236
 background for, 225–226
 data, 229–230
 demographic neighbors, 229
 in-sample fit, 231
 MCMC inference, 228
 model for, 226–228
 modeling of, 225–235
 out-of-sample fit, 231
 prior specification, 228
 R implementation, 235
 results of, 230–235
 sample statistics, 230*t*
 zip code, sample size for, 231*f*
- Interior solution, 272
- Invariant distribution, 60
- Irreducible chain, 60
- Joint distribution, 10
- Kuhn–Tucker first-order condition, 126
- Label-switching problem, 81
 phenomenon of, 83
- Λ
 posterior distribution of, 249, 249*f*
- Laplace approximation, 36
- Latent preference
 unconditional expectation of, 226
- Latent variable
 correlation with binary logit parameters,
 112*f*

- formulation of, 104
 - multivariate, 105
- Latent variable models, 104–106
 - ordered models, 104
- Law of large numbers, 51
- Lebesgue measure, 14, 42
- Likelihood function, 270
- Likelihood principle, 15
- Linear hierarchical models, 73*f*
 - shrinkage in, 74*f*
- Linear models
 - hierarchical, 70–75
- Linear utility function, 122
- Linear utility model, 269
- Lists, 287–291
- Log-likelihood function
 - Hessian of, 19
- Loops, 294–295
 - implicit, 295
- Loss function, 17
 - nonlinear, 17
- LU decomposition, 24

- Marginal density, 11
- Marginal utility, 270
 - constant, 123
 - deterministic part of, 210
- Marketing
 - Bayesian literature in, 159
- Marketing decisions, 177–180
 - plug-in, 179
 - vs full Bayes approach, 179
- Marketing mix variables, 271
 - non-random, 200–206
 - framework for, 200–202
- Marketing practice
 - findings on, 156–158
 - influence on, 156–158
- Marketing problems
 - basic paradigm for, 2
 - Bayesian approach
 - benefits of, 4–6
 - costs of, 4–6
 - case studies
 - overview of, 6
 - Cholesky root, 6
 - example of, 3–4
 - methodological material
 - overview of, 6
 - multinomial logit model, 3
 - multivariate normal density, 3
 - sales response model, 185
 - simultaneity, 6
 - statistical modeling
 - components of, 2
 - units, 2
 - unit-level behavior, 2
 - within-unit model, 3
- Markov chain, 50, 255
 - burn-in period, 51
 - continuous state space, 58
 - discrete approximation, 58*f*
 - discrete state space, 58
 - double exponential density, 58*f*
 - ergodic, 51
 - irreducibility, 63
 - irreducible, 60
 - kernel, 62
 - posterior expectation, 50
 - reducible, 60
 - reversibility principle, 62
 - stationary distribution of, 50, 62
- Markov chain Monte Carlo (MCMC)
 - methods, 1, 49–102
 - conditional distributions, 67–70
 - data augmentation, 75–79
 - diagnostics of, 99–102
 - directed graphs, 67–70
 - Gibbs sampler
 - bivariate normal, 52–57
 - hierarchical linear models, 70–75
 - hybrid, 97–99
 - integration estimate, 51
 - metropolis
 - algorithm of, 86–94
 - micro-econometric problems, 49
 - normal mixtures
 - identification in, 81–84
 - normals model
 - mixtures of, 79–85
 - output of, 327
 - performance of, 100
 - probit example, 75–79
 - sampling theory, 51
 - tips on, 327
- Markov chain Monte Carlo (MCMC) model
 - Metropolis–Hasting approach, 201
- Markov chain sampler, 86
- Markov chain theory, 57–63
 - time reversibility, 60
- Matrices, 288–291
- Matrix operations, 295–296
- Maximum likelihood estimator (MLE), 18
- Memory, 7

- Metropolis
 - algorithm of, 86–94, 134
 - continuous, 87
 - convergence, 89
 - discrete, 87
 - continuous state space, 88
 - kernel, 88
 - transition function of, 88
- Metropolis chain, 88
 - independence, 89–90
 - random-walk, 90–92
 - Gaussian, 91
 - optical scaling of, 92
 - scaling of, 92–94
- Metropolis within Gibbs, 98
- Metropolis–Hastings algorithm
 - random-walk, 314
- Miller Lite, 220
 - cumulative distribution of, 223*f*
- Mixture components
 - identification of number of, 148–149
 - Schwarz approximation, 148
- Model fit, 231
 - comparison of, 232*t*
- Monte Carlo integration, 37
- Multinomial distribution, 44–45, 105
- Multinomial logit (MNL) model
 - hierarchical, 136–142
 - independence metropolis
 - vs random-walk metropolis, 97*f*
 - metropolis
 - algorithm of, 94–97
 - optical scaling of, 96*f*
 - unit-level, 129
- Multinomial probability
 - computation of, 240*f*
- Multinomial probit (MNP) model, 106–116, 159
 - diagonal, 173–177
 - diagonal vs identity, 175*f*
 - Gibbs sampler, 108, 111*f*
 - autocorrelation properties of, 110–113
 - likelihood for, 114–116
 - regions of integration for, 115*f*
- Multiple regression, 21–26
- Multiple unit purchases
 - frequency of corner for, 273*t*
 - frequency of interior solutions for, 273*t*
- Multivariate analysis
 - example of, 21–35
- Multivariate normal density, 80, 146
- Multivariate normal distribution, 42–43
 - evaluation of integrals
 - GHK method for, 39–40
 - finite mixtures of, 79
- Multivariate probit model, 116–121
 - correlations, 120*f*
 - Gibbs sampler, 118
 - intercepts, 119*f*
 - Scotch survey data, 119
 - with improper prior on Σ , 121
- Multivariate regression, 31–34
 - $\beta \mid \Sigma$, 47
 - Sigma*, draw of, 46
 - simulation from posterior of, 45–47
 - model of, 45
 - prior, 45
- Multivariate regression model (MRM), 65
- Natural conjugate prior, 23
- Newton–Raftery (NR) estimate, 168
- Nonhomothetic choice model, 123–124
- Nonidentified (NID) sampler, 109
 - vs ID MNP samplers, 114*f*
- Normal mixtures
 - artificial identification constraints, 82
 - density draw of, 83*f*
 - finite-mixture model, 144
 - five-component, 151*f*
 - Gibbs sampler, 145–148
 - hierarchical models
 - applications of, 149–153
 - identification in, 81–84
 - label-switching for, 82*f*
 - multivariate
 - frequency of, 85*f*
 - marginal densities of, 86*f*
 - one-component, 151*f*
 - use of, 142–153
- Normals model
 - mixtures of, 79–85
- Numeric loop index, 294
- Out-of-sample fit, 231
- Outcome variable
 - multinomial, 106
- Packaged goods, 219–224
 - background for, 207–209
 - choice model for, 207–224
 - convex budget set, 214*f*
 - data description of, 215*t*
 - data for, 214–219

- discrete quantity dealing, 207–224
- dominating linear set, 214*f*
- first-order conditions, 209*f*
- likelihood specification for, 211–213
- model for, 209–214
- piecewise linear budget set, 214*f*
- price points, 223
- quantity discounts, 207–224
- R implementation of, 224
- solution procedure for, 211–213
- utility minimization, 209*f*
- Panel data structures, 133
- Parameter estimation, 17
- Posterior distribution, 14
- Posterior model probabilities, 159
 - with unidentified parameters, 170–171
- Posterior odds ratio, 160
- Predictive distribution, 16
- Price coefficient
 - posterior distribution of, 140*f*
- Price points, 223
- Price schedule
 - utility maximization, 210*f*
- Pricing policy
 - implications for, 275–277
- Probability
 - non-normality of, 78*f*
- Probability mass function, 10
- Probit model
 - part-worth estimates for, 261*t*
- Program editing, 284
- Pseudo-random number generator, 41
 - univariate, 41
- Purchase quantity
 - distribution of, 272*t*

- Quantile function, 42
- Quantity discounts, 207

- R environment
 - setting up of, 279–284
- R implementation, 277
- R language, 7, 285–303
 - accessing elements, 291–293
 - built-in functions, 286–287, 296
 - debugging functions, 298–299
 - graphics, 299–301
 - histogram labelling, 300*f*
 - multiple plots, 302*f*
 - inspecting objects, 287–288
 - lists, 288–291
 - subsetting, 291–293
 - loops, 294–295
 - implicit, 295
 - matrix, 288–291
 - subsetting, 291–293
 - matrix operations, 295–296
 - obtaining help in, 280–283
 - obtaining of, 279–280
 - obtaining parameter estimates, 315–321
 - program editing, 284
 - using R GUI, 280, 281*f*
 - shortcut
 - customizing of, 280
 - simple debugger, 299
 - system information, 301–303
 - timing, 303
 - user-defined functions, 296–298
 - vectors, 288–291
 - subsetting, 291–293
 - workspace, 287–288
- R package, 279, 323
- Random effects model, 133
- Random permutation, 84
- Random utility error, 211
- Random walk chains
 - comparison of, 139*f*
- Random-walk metropolis chain, 90–92
 - Gaussian, 91
 - optical scaling of, 92
 - scaling of, 92–94
- Rate function, *See* Density function
- Recursive system, 186
- Reducible chains, 60
- Regression, 21–35
 - assessing priors for, 26–28
 - coefficients, 26
 - error variance, 70
 - LU decomposition, 24
 - multiple, 21–26
 - natural conjugate prior, 23
 - noninformative prior, 28
 - priors, 30–31
 - standard linear model, 22
 - univariate, 33
 - variance–covariance matrix, 22
- Retail scanner data, 103
- Reverse chain
 - transition matrix of, 61
- RW Metropolis draw, 271

- SAMP, 203
- Sampler
 relative numerical efficiency, 92
- Savage–Dickey density ratio, 162
- Scale usage heterogeneity, 237, 239*f*,
 247–249
 respondent range vs median, 248*f*
- Scanner data, 130
 store-level, 158
- Scanner-panel data, 271
- Schwarz approximation, 148
- Schwarz criterion, *See* Bayesian information
 criterion (BIC)
- Seemingly unrelated regression (SUR) model,
 34, 65–67
- Selected variance parameter
 posterior distribution of, 141*f*
- Shrinkage
 adaptive, 132
- σ_i
 prior distributions for, 243*f*
- Sigma*
 LU decomposition, 47
 QR decomposition, 46
 UL decomposition, 47
- Simple debugger, 299
- Simultaneity, 185–206
 structural models of, 195–200
- Standard deviation (STD), 37
- Standard error, 37
 numerical, 38
- Statistical inference
 goal of, 13–15
- Stock-keeping unit (SKU), 208
- Student t distribution, 42–43
- Subutility function, 125
- Supply model
 profit-maximizing prices, 197–198
- System information, 301–303
- Time reversibility, 60
- Truncated distributions, 41–42
- Uniform distribution, 11
- Unit-level, 103
 models, 103–128, 182
- Univariate distribution, 108
- Univariate regression, 33
- Unknown quantity, 13
- User-defined functions, 296–298
- V_β
 posterior mean of, 322*t*
- Variety
 demand for, 125–128
 modeling consumer demand for,
 269–277
 background of, 269–270
 compensating values, 274–275
 data, 271
 model of, 270–271
 results of, 273
- Vectors, 288–291
- Wishart density
 inverted, 29
- Wishart distribution, 30–31, 43–44
 inverted, 29, 43–44
- Wishart prior
 inverted, 66, 189
- Workhorse model, 35
- Yogurt data
 frequency of corner and interior solutions in,
 272*t*
 parameter estimates, 274*t*